

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2002, 04:14:19 ; Search time 3106.14 Seconds  
(without alignments)  
23342.510 Million cell updates/sec

Title: US-09-652-292-1  
Perfect score: 4395  
Sequence: 1 gaggggtcttgcagcc.....attatttgtaaaaaaaaaa 4395

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_om:\*
- 20: em\_or:\*
- 21: em\_ov:\*
- 22: em\_pat:\*
- 23: em\_ph:\*
- 24: em\_pl:\*
- 25: em\_ro:\*
- 26: em\_sts:\*
- 27: em\_sy:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htgo\_hum:\*
- 31: em\_htgo\_inv:\*
- 32: em\_htgo\_rod:\*
- 33: em\_htg\_hum:\*
- 34: em\_htg\_inv:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4395	100.0	4396	9	AF248053	AF248053 Homo sapi
2	2465	56.1	4167	9	AF321240	AF321240 Homo sapi
3	2424	55.2	4126	9	HS28H201	AL137188 Novel hum
4	1810	41.2	127418	9	HS28H20	AL031055 Human DNA
C 5	182	4.1	385	6	AX150151	AX150151 Sequence
C 6	156	3.5	411	11	G31518	G31518 human STS S
C 7	99	2.3	1520	9	BC010033	BC010033 Homo sapi
C 8	99	2.3	1524	9	BC005060	BC005060 Homo sapi
C 9	99	2.3	17559	2	AC009133	AC009133 Homo sapi
C 10	99	2.3	181705	2	AC023831	AC023831 Homo sapi
C 11	67	1.5	70876	9	AC023105	AC023105 Homo sapi
C 12	64	1.5	185926	2	AC027632	AC027632 Homo sapi
C 13	63	1.4	108162	2	AC020960	AC020960 Mus muscu
C 14	63	1.4	113982	2	AF192303	AF192303 Homo sapi
C 15	63	1.4	122772	2	AC008874	AC008874 Homo sapi
C 16	63	1.4	164774	2	AC019104	AC019104 Homo sapi
C 17	63	1.4	169241	2	AL445987	AL445987 Homo sapi
C 18	63	1.4	176232	2	AL354659	AL354659 Homo sapi
C 19	63	1.4	198027	2	AC008515	AC008515 Homo sapi
C 20	63	1.4	298734	2	AC008972	AC008972 Homo sapi
C 21	62	1.4	3351	9	AF048726	AF048726 Homo sapi
C 22	62	1.4	75778	9	AL160260	AL160260 Human DNA
C 23	62	1.4	150009	2	AC022282	AC022282 Homo sapi
C 24	62	1.4	150856	9	AC087072	AC087072 Homo sapi
C 25	62	1.4	154611	2	AC068893	AC068893 Homo sapi
C 26	62	1.4	166872	2	AC021033	AC021033 Homo sapi
C 27	62	1.4	169725	2	AC073367	AC073367 Homo sapi
C 28	62	1.4	183029	2	AC021030	AC021030 Homo sapi
C 29	62	1.4	186581	2	AC021031	AC021031 Homo sapi
C 30	62	1.4	193877	2	AC009686	AC009686 Homo sapi
C 31	62	1.4	195170	2	AL353576	AL353576 Homo sapi
C 32	62	1.4	196822	2	AC021041	AC021041 Homo sapi
C 33	62	1.4	196933	2	AC079854	AC079854 Homo sapi
C 34	62	1.4	197440	2	AC068580	AC068580 Homo sapi
C 35	62	1.4	208561	2	AC069540	AC069540 Homo sapi
C 36	61	1.4	63545	2	AC087639	AC087639 Homo sapi
C 37	61	1.4	161204	2	AC025772	AC025772 Homo sapi
C 38	61	1.4	166211	2	AC079135	AC079135 Homo sapi
C 39	61	1.4	166263	2	AC019047	AC019047 Homo sapi
C 40	61	1.4	177112	2	AC007498	AC007498 Homo sapi
C 41	61	1.4	185272	2	AC090734	AC090734 Homo sapi
C 42	61	1.4	189036	2	AC019142	AC019142 Homo sapi
C 43	61	1.4	245081	2	AC074138	AC074138 Homo sapi
C 44	60	1.4	85173	9	AL449143	AL449143 Human DNA
C 45	60	1.4	98404	2	AL355676	AL355676 Homo sapi

ALIGNMENTS

RESULT 1

AF248053  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

AF248053 4396 bp mRNA 12-APR-2001  
Homo sapiens glucose transporter (GLUT10) mRNA, complete cds.  
AF248053  
AF248053.1 GI:13603726  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 4396)  
Fossey,S.C., Mihic,S.J., Craddock,A.L., Mychaleckyj,J.C.,  
Dawson,P.A. and Bowden,D.W.  
GLUT10: A novel glucose transporter in the type 2 diabetes linked  
region of chromosome 20q12-13.1  
Unpublished  
2 (bases 1 to 4396)  
Fossey,S.C., Mihic,S.J., Craddock,A.L., Mychaleckyj,J.C.,  
Dawson,P.A. and Bowden,D.W.

Query Match	100.0%	Score 4395;	DB 9;	Length 4396;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 4395;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	gaggggctcttgcagagcctgagggcgggcgggcgggcggtcctgggtccctccgtcccg	60	
Db	2	GAGGGGGTCTTGCCAGGCTGGGGCGCGGGGGGGTCTGGGGTCCCTCCCTCGTCCCG	61	
QY	61	ctccaggcctcgggacctggctggccgacgtggcgttgccgagcgtgcgcggagag	120	
Db	62	CTTCCAGGCGCTGGGGCCTGGCTGGCCGACGTGGCGCTGGCGGCGTGGCGGGAGGG	121	
QY	121	cagggcaggagagacagagcgagggcgggcggaagtgttcgcgcgcgcgcgcgtt	180	
Db	122	CAGGCGCAGAGGGACAGAGCGGGCGGGCGGAAAGTTTGTCCGGCGGCGAGCGCGTT	181	
QY	181	ggggactcggcgggggaatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	240	
Db	182	GGGGACATCCGGGGGGGATGCGCGCGCGGCCCTCAGCGCCCCAGCAGCGCCGAGT	241	
QY	241	ccgctcgcacatgggcacatccccacatcctgccttctgtgcctctgtcttctgtc	300	
Db	242	CCGCTCGCATGGGCCATCCCGACATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTG	301	
QY	301	ggtgcctgaaccttgggttatgaactggcagtcacatcaggtgcacctgcgcacgtcag	360	
Db	302	GGTGCCCTGACCTTGGTTATGAATGGCAGTCAATCAGTGCCTGCTGCTGCCACTGC	361	
QY	361	cttgaccttgggctaagctgcttggagcaggaagtcctggtaggcagcctgcctcggg	420	
Db	362	CTTGACTTTGGCTAAGCTGCTTGGAGCAGGAGTCTCTGGTGGGCGAGCTGCTCT	421	
QY	421	gctcctcgcctcctcctggttgggtgctcctcctcctcctcctcctcctcctcctc	480	
Db	422	GCTCTCTCGGCTCCCTGGTTGGTGGCTCTCTCATTTACTGCTATGGCAGGAAGCA	481	
QY	481	atcctcgggagcaacttggctgctgcgcagcagcctgcacctaggcctaggctgtcc	540	
Db	482	ATCTTCGGAGCACTTGGTGCTGCTGCGAGGACCTGACCCCTGGGCGCTGCTGT	541	
QY	541	ctggcctgctggctcctggcgcgctgtgtgtggcttcgcattcctcctcctcctc	600	

Qy 1681 accttccctgctctacaggactgaccgctgtctctcgccctgggcttccatctctatttattgtt 1740  
Db 1682 accttccctgctctacaggactgaccgctgtctctcgccctgggcttccatctctatttattgtt 1741  
Qy 1741 cctgaacaaaggccctgctgagagatagaccagcagcttccaaagagagcttc 1800  
Db 1742 cctgaacaaaggccctgctgagagatagaccagcagcttccaaagagagcttc 1801  
Qy 1801 acctgagcttggccacaggcagaaactccactggcctcccgatcccgatcagccgacatcagatc 1860  
Db 1802 acctgagcttggccacaggcagaaactccactggcctcccgatcccgatcagccgacatcagatc 1861  
Qy 1861 tctgggctcctctgaggaatccgctgctggaaactctggaaactctgggactctgggagct 1920  
Db 1862 tctgggctcctctgaggaatccgctgctggaaactctggaaactctgggactctgggagct 1921  
Qy 1921 atctccagcactcctctcctagggcccgagacacagctccagctgctgctctcttgggagct 1980  
Db 1922 atctccagcactcctctcctagggcccgagacacagctccagctgctgctctcttgggagct 1981  
Qy 1981 ggcctctgcccacaaagctggttcttctgctggggttaaaagagatgaaagcttggagaa 2040  
Db 1982 ggcctctgcccacaaagctggttcttctgctggggttaaaagagatgaaagcttggagaa 2041  
Qy 2041 tggccaaatcttctcatttgggttctcaggccctgagggcttcttgagagatcagtttcatgctc 2100  
Db 2042 tggccaaatcttctcatttgggttctcaggccctgagggcttcttgagagatcagtttcatgctc 2101  
Qy 2101 ctggttctcccatgacttgacatcttctgagcttcttctatagagagatctctatgaa 2160  
Db 2102 ctggttctcccatgacttgacatcttctgagcttcttctatagagagatctctatgaa 2161  
Qy 2161 gctcttctggcccatgacttcttctgaggggttaaaagatctcaggggtaccaaactccgggagag 2220  
Db 2162 gctcttctggcccatgacttcttctgaggggttaaaagatctcaggggtaccaaactccgggagag 2221  
Qy 2221 gttttcccgatatacaccctaaataccaaatgagagatcattcttctctaaactctctt 2280  
Db 2222 gttttcccgatatacaccctaaataccaaatgagagatcattcttctctaaactctctt 2281  
Qy 2281 tcaactgctgggacatcttctgaggggttaaaagatcttcttcttctctctctctctct 2340  
Db 2282 tcaactgctgggacatcttctgaggggttaaaagatcttcttcttctctctctctctct 2341  
Qy 2341 ttttttggggtgagctctcattctgttggccaggtggtgctgacttctggtcactgcaac 2400  
Db 2342 ttttttggggtgagctctcattctgttggccaggtggtgctgacttctggtcactgcaac 2401  
Qy 2401 ctccactctcctgggttcaagcgatctcctgctcagcctcctcctcagtagctggattaca 2460  
Db 2402 ctccactctcctgggttcaagcgatctcctgctcagcctcctcctcagtagctggattaca 2461  
Qy 2461 ggcggtgcccacacacccagcctaatttatttttagcagagatggggttctcactgtgtg 2520  
Db 2462 ggcggtgcccacacacccagcctaatttatttttagcagagatggggttctcactgtgtg 2521  
Qy 2521 gcccagctgggtgagctcctgagctcagtgatccacccacccctcagcctcccgagctg 2580  
Db 2522 gccagctgggtgagctcctgagctcagtgatccacccacccctcagcctcccgagctg 2581  
Qy 2581 ctaggattacaggccttcttgactcttcttctctgagtttcttcttcttcttcttcttctt 2640  
Db 2582 ctaggattacaggccttcttctgactcttcttctctgagtttcttcttcttcttcttctt 2641  
Qy 2641 acccagaatatttaccctccacacacactgactcttcttcttcttcttcttcttcttctt 2700  
Db 2642 acccagaatatttaccctccacacacactgactcttcttcttcttcttcttcttcttctt 2701  
Qy 2701 gtcctgtctgctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2760  
Db 2702 gtcctgtctgctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2761

Qy 2761 gtataatgggaagccctgtaccagggtcattctttaagattttctctgactccagtgagctgg 2820  
Db 2762 gtataatgggaagccctgtaccagggtcattctttaagattttctctgactccagtgagctgg 2821  
Qy 2821 aatctcaaatgctggtcttagggactgtctccaggaagggtgaggaaggcttcttcttcttctt 2880  
Db 2822 aatctcaaatgctggtcttagggactgtctccaggaagggtgaggaaggcttcttcttcttctt 2881  
Qy 2881 gagatgggttggagggcccaaacctgcttctcaatatattgccttcttcttcttcttcttctt 2940  
Db 2882 gagatgggttggagggcccaaacctgcttctcaatatattgccttcttcttcttcttcttctt 2941  
Qy 2941 cttgaaacttgaglaaataaacaactccctgaacctgaatttctctcaatctgcagaaatgggga 3000  
Db 2942 cttgaaacttgaglaaataaacaactccctgaacctgaatttctctcaatctgcagaaatgggga 3001  
Qy 3001 taattatgtccaggggttatatttagacctgttctcttccagggaggttccccagctgggt 3060  
Db 3002 taattatgtccaggggttatatttagacctgttctcttccagggaggttccccagctgggt 3061  
Qy 3061 ccaggccctgggaatttctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3120  
Db 3062 ccaggccctgggaatttctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3121  
Qy 3121 aagggtcaggggtgaatcagatggggactgagcaagtagctatgactgagatcagataa 3180  
Db 3122 aagggtcaggggtgaatcagatggggactgagcaagtagctatgactgagatcagataa 3181  
Qy 3181 gaaaggactgacaaagcctccagatgctggggagaaaggagagcctaaagaaagaaagaa 3240  
Db 3182 gaaaggactgacaaagcctccagatgctggggagaaaggagagcctaaagaaagaaagaa 3241  
Qy 3241 aggtgctgagctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3300  
Db 3242 aggtgctgagctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3301  
Qy 3301 acctggtcctcagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3360  
Db 3302 acctggtcctcagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3361  
Qy 3361 agtgcctggagctggcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3420  
Db 3362 agtgcctggagctggcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3421  
Qy 3421 tacaagctggtaaacacagccattataaaataaaatgataaaatgataaaatgataaaatgata 3480  
Db 3422 tacaagctggtaaacacagccattataaaataaaatgataaaatgataaaatgataaaatgata 3481  
Qy 3481 aattacataaaacaaacaaatataactcaaaatctacttaattcttcttcttcttcttcttctt 3540  
Db 3482 aattacataaaacaaacaaatataactcaaaatctacttaattcttcttcttcttcttcttctt 3541  
Qy 3541 ctattatctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3600  
Db 3542 ctattatctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3601  
Qy 3601 acaccgcgactccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3660  
Db 3602 acaccgcgactccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3661  
Qy 3661 aactgctggtgggagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3720  
Db 3662 aactgctggtgggagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3721  
Qy 3721 tgaattatgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3780  
Db 3722 tgaattatgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3781  
Qy 3781 caaattaaactttaaagctatgcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3840  
Db 3782 caaattaaactttaaagctatgcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3841  
Qy 3841 atggagggtatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3900





Db 2881 TTCTACTTATCTCATTACCCAGGCTCCCTCTTTGGACCTGTAAAGGGTCAGGGTCAAT 2940  
 QY 3137 cagatggggactgagcaaatgaactatgactgagatcatgataagaaagggactgacaag 3196  
 Db 2941 CAGATGGGGACTGAGCAAGTACTATGACTGCAGATCATGTAAAGAAAGGACTGACAAG 3000  
 QY 3197 aagctccacagatgctggggagaaagaagagctaaaaatagatccctagctgctgctgctt 3256  
 Db 3001 AAGCTCCACAGATGCTGGGAGAAATGAAGAGCTAAATAGATCTAGTCTGGATCCTTT 3060  
 QY 3257 gtccatccatgctgacacatatgggtgctggcagagcccccaaggagactctggcctctcgag 3316  
 Db 3061 GTATCCATCGTGCACATATGGGTGCTGCGAGAGCCCCAAGGACTCTGGCCTCTCGAG 3120  
 QY 3317 ttctccctattctccattctagatgcttcccttctgattccctgattccagtgatgctggagctgac 3376  
 Db 3121 TTCTCTATCTCTCCATCTAGATGCTTCCCTTGTATCCAGTATGCTGAGAGTGGC 3180  
 QY 3377 ttgccaagctgtgagagctggtgctacattttcagagattttcacaagttggtaaacaa 3436  
 Db 3181 TTTGCCAAGCTGTGAGAGCTGCTTACATTTTCAGGATTTTACAGTTTGGTAAACA 3240  
 QY 3437 cagccattataaaaaattaaatgatttaattataattataattacattacattacataaaacaa 3496  
 Db 3241 CAGCCATTATAAAAAATTAAATGATTAAATTTAATTAATTAAGTAAATACATTAACAA 3300  
 QY 3497 aaaaattactcaaaattcattacttaatttactcactgttactatttactatttctgctt 3556  
 Db 3301 AAAAAATTACTCAAAATTCATTACTTAATTTACTACCTGTCTATTATCTTGTGCTTT 3360  
 QY 3557 tgaagctatttctacataagtaactcttattgagacactaggggagacacgcgcactctctt 3616  
 Db 3361 TGAGGCTATTTTACATAGTAATCTCTTATGGAGACTAGGGGAGACACCGGCATCTCTT 3420  
 QY 3617 cctgattccccactcaatgacatcatgcttagcttctggtgcttcaactgctgctggag 3676  
 Db 3421 CCGTATTCCTCCACTCAATGACATCATGTTAGTCTTTGTTGCTTAACTGGCTGTGGGAG 3480  
 QY 3677 ttttttctacaaaagattagagagactacacatcagggctgatttatttctt 3736  
 Db 3481 TGTTTTGTGATCACAAAGATTAGAGAGGACTACATCAGGGCTTGATTTATTTGTTGTT 3540  
 QY 3737 gatttctagacttcagaaactgctgataaaatgctgaatgataatgataaatttaacattaaa 3796  
 Db 3541 GATTTTCTAGACTTCAGAACATGCTGGATAAATGTCAGTAAATGCAAAATTAACCTTTAAA 3600  
 QY 3797 gctgctcttctgttagcccaatatacagtgtagcaccacaaaatggagggtattctt 3856  
 Db 3601 GTATGCTTGTGTGAGCCAATACATGCTGTATAGCACCAAAAATGGAGGATTTCT 3660  
 QY 3857 tccagtagtgaaactgtccatccgtttcagctgacagctgctcaaatcatttaagaagg 3916  
 Db 3661 TCCAGTAGTTGAACACTGTCATCCGTTTCAGCTGACAGCTGCTCAAAATCAATTAAGAAGG 3720  
 QY 3917 agttctgacattcattcttctgcttcttcttcttcttcttcttcttcttcttcttcttctt 3976  
 Db 3721 AGTTCTGACATCAATTTTCATGTTTACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3780  
 QY 3977 tccaaccagcattcagccgaacccataaccattcttcagtgccctgactgacagttatc 4036  
 Db 3781 TTCAACAGCATTCATGCGGAACCTATACCATCTTCTTCTGCTGCTGCTGCTGCTGCTGCT 3840  
 QY 4037 aggaatttttctgtagtctcaatttttctcaaatcattgcccgaatcgcagtgatgtga 4096  
 Db 3841 AGGGATTTTATTTGTAGCTTAATTTTGTCAAAATCATGGCCAAATCGCAGTATAGTTGA 3900  
 QY 4097 ctttggatcacaaggttttggcaaaaaaaataatacaaaaaattcttctgtaagaatcaat 4156  
 Db 3901 CTTTGGATACAAGCTTTGGCAAAAAAATAATTAACAATAATTTCTGTAAGAATCAAT 3960  
 QY 4157 tggctatatggaatttgggataagaataatttacaataaagaataatttacaataaagaagt 4216  
 Db 3961 TGCTATATGGAATTTAGGATAAAGAATATTTACAATAAAGAATATTTACAATAAAGAGT 4020

QY 4217 ttattattattgttaagtgtgtgcaacaaacatacccttttatctctgtataaatttatac 4276  
 Db 4021 TTATTATTATTGTGAAGTTGTGTGCAACAAACATACCTTTTATCTCTGTAAATTTATAC 4080  
 QY 4277 acacaaaaatttaacaaagattctgtgaagaataatttggtctattggaatttagataga 4336  
 Db 4081 ACACAAAAAATTAAACAAAGATTCTGTAGAAATTAATGGCTATATGGAATTTAGGATAGA 4140  
 QY 4337 atattacaataaagagattttacaat 4363  
 Db 4141 ATATTACATAAAGAGTATTTACAAT 4167

RESULT 3  
 HS28H201  
 LOCUS HS28H201 4126 bp mRNA PRI 21-NOV-2000  
 DEFINITION Novel human gene mapping to chromosome 20, similar to membrane transporters.  
 ACCESSION AL137188  
 VERSION AL137188.3 GI:11322734  
 KEYWORDS human.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 4126)  
 AUTHORS Stavrides,G.S., Hashim,Y., Huckle,E.J. and Deloukas,P.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-JAN-2000) E-mail contact: hunquery@sanger.ac.uk  
 COMMENT On Nov 23, 2000 this sequence version replaced gi:11065679.  
 This cDNA sequence was assembled from public domain ESTs and single pass sequencing reads from expressed DNA templates, aligned to the genomic DNA sequence from the bacterial clone 28H20 (AL031055). The EST sequences listed match this sequence with an identity of at least 95% between the coordinates shown.  
 Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr20/ Sanger Centre name : d128H20.C20.1  
 Location/Qualifiers  
 1. .4126  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20q"  
 1. .112  
 /number=1  
 9. .1634  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA569822.2"  
 /db\_xref="GI:11065680"  
 /translation="MGHSPVLPCLASVSLGGTLFGYELAVISGALLPLQDFGLSC LQEFIVGSLLLGALLVGGFLIDCYGRKQAILGSLNVLVLAGSLTGLAGSLAWLV LGRVYVGTATSLSSMACCIYVSELVGRQGVLYEAGITVGILLSYALAVLAGT PWCWRHMFQWATAPAVLIQSLILFLPAGTDETHKDLIPLOGGERPKLGRPRPYSF LIDLFRADNNRGRTTTCGLGLVLFQOLTGPVNLVYASTIFSSVGFHGGSSAVLSVGL GAVYCAATLTAMGLVDRAGRRALLAGCALMALSVSGILGVFAVPMDSGSPCLAVPN ATGCTPGDGLQDSSLPPIPTNEDQREPIILSTAKTKPHPRSGDPSAPRLALS CNSFWAANLFISSLFDLIGTIGLSWTFELLYGLTAVLGLGFIYLPVETKGSGLAEI DQFQKRRTLSFGHRQNSTGIPYSRIEISAAS"  
 13. .1296  
 /number=2  
 1297. .1419  
 /number=3  
 1338. .1643  
 /note="matches EST AA313045"  
 1420. .1555  
 /number=4  
 1556. .4126  
 /number=5  
 2095. .2137

FEATURES  
 source  
 exon  
 CDS  
 exon  
 misc\_feature  
 exon  
 exon  
 misc\_feature





Db	601	TCATCCCACTCCAGCGAGGTGAGCGCCCCCAAGCTGGCGCCCGGGAGGCCACAGGTACTCCT	660
QY	902	ttctggaactcttcaggagcagcgataaactgtagagccggaccacagtggggccttggggc	961
Db	661	TTCTTGGAGCTCTTCAGGGCAGCGATAACTATGTCAGGCGCGGACCACTAGTGGGCTTGGGGC	720
QY	962	tgggtctcttcagcaactaacagggcagcccaaacctgctgtgctatgcctccaccatct	1021
Db	721	TGGTGCTCTTCCAGCAACTAACAGGCGAGCCCAAGCTGTGTCTATGCTCTCCACCACT	780
QY	1022	tcagctcgttgggtttccatg9999gatccctcagccgtctggtgcccctctgtggggcttggcg	1081
Db	781	TCAGCTCCGTTGGTTTTCATGGGGATCTCTACGCCGTGCTGGCCCTCTGTGGGGCTTGGCG	840
QY	1082	cagtgaaagttgagcgtaccctcagccgcatagggcctggtgagccgtgagccgcaggg	1141
Db	841	CAGTGAAGTGGCAGCTTACCCTCAGCGCCATGGGGCTGTGACCGTGTGACGCGCGCAGGG	900
QY	1142	ctctgtgtcagctggctgtgcctcatagggccctgtccgtcagtgagcataagcctctca	1201
Db	901	CTCTGTTGCTAGCTGGCTGTGCCCTCATGGCCCTGTCTGCTCAGTGGCATAGGCTCTGTCA	960
QY	1202	gctttgcccgtgccatggactcagggcccaagctgtctggctgtgcccactgccaccgggc	1261
Db	961	GCTTTGGCGTCCCATGGACTCAGGCCCAAGCTGTCTGGCTGTGCCCAATGCCACCGGGC	1020
QY	1262	agacaggcctccctggagactctggcctgtcagagactcctctatcctcccatcccaa	1321
Db	1021	AGACAGCCTCCCTGGAGACTCTGGGCTCTGTCAGGACTCTCTCTTACTTCCCATTCCAA	1080
QY	1322	ggaccaatgaggaccaaaggagccaaactctgtccactgtctaaagaaaccaagcccatc	1381
Db	1081	GGACCAATGAGGACCAAGAGGCCAATCTTGTCCACTGCTTAAGAAACCAAGCCCCCTC	1140
QY	1382	ccagatctggagacccctcagccctctcggctcggccctgagctctgcctccctgggc	1441
Db	1141	CCAGATCTGAGAGCCCTCAGCCCTCTCTCGGCTCTGGCTGGCCCTGAGCTCTGCCCTGGGC	1200
QY	1442	ccctctgcccctcgggggcactgactgctgagcgtgagacgcaactgctgctcatgg	1501
Db	1201	CCCTCTGCCCGTCTGGGGGATGACACTGTGCGCTGTGACCGCACTGCTGTGCTGATGG	1260
QY	1502	tccttgcagtgcccttctccttgggttggggcagtgacctgctgtcctcagcgaga	1561
Db	1261	TCCTTTGTCTAGTGCCTTCTCTTTGGTTTGGCCAGTCACTGGCTGTCTCAGCGAGA	1320
QY	1562	ttctacccttggaatacagagaagacgttgcctctcgaacagcttcaactggcg	1621
Db	1321	TCATCCCTGTGGAGATACGAGAAGAGGCTTTCGGCTTCTGCAACAGCTTCAACTGGGGCG	1380
QY	1622	ccaacctcttcacagcctcctctccctcgatctcatctgacacatcggctgtccatga	1681
Db	1381	CCAACTCTTCATCAGCCTCTCTTCTCGATCTATTTGGACCATCGGCTTGTCTCGA	1440
QY	1682	ccttcctgtctacggactgaccgctgtccctcggcctgggcttcactatatttattgttc	1741
Db	1441	CTTCTCTGCTCTACGACTACCGCTGTCTCTCGGCTTCCCTGGCTTCACTATTTATTTGTTTC	1500
QY	1742	ctgaacaagagccagtcgttggcagagatagaccagcaggttccagaagacaggttca	1801
Db	1501	CTGAACAAGAGGCCAGTCTGTTGGCAGAGATAGACCAAGCATTTCCAGAAGAGACGCTTCA	1560
QY	1802	ccttgagcttggccacagggcagaactccactggcatccctgacagccgcatcagatct	1861
Db	1561	CCCTGAGCTTTGGCCACAGGCAGAACTCCACTGGCATCCGTTACACCCGCTCTGAGACT	1620
QY	1862	ctggggcctcctgaggaatccgtgtcctgg-aattctggaactgtggccttggcagacc	1920
Db	1621	CTGCGGCTCTCTGAGGAATCCGCTGCTCGCTGGAAATCTTGGAACTGTGGCTTTGGCAGACC	1680
QY	1921	atctcagcatcctgtctcttagcccccagagcacagttccagcttggtcttttggaggt	1980
Db	1681	ATCTCCAGCATCTCTGCTTCTTAGGCCCCAGACACAAGTTCCAGCTGGTCTTTTGGGAGT	1740

Qy	1981	ggccccgccccaaagggtttgcttttggctgggttaaaagatgaagtttgaaa	2040
Db	1741	ggcccttgcctcccaaaagggtgctgttctggctgggttaaaagatgaagtttgaaa	1800
Qy	2041	tgcccaattcttcatttgggtttcaggccctgaaggtctcttggagatctagttcattgc	2100
Db	1801	tggcccaactcttctatttttgagcttgcaggccctgaaggtctctcaggatcttagcttcatgc	1860
Qy	2101	ctcgggtttccccattgacttggacatttttgcagtttttataagaagaatatcttatgaa	2160
Db	1861	ctcagttttcccatctgacttggacatctctgcagctattttatagacgaattatttctatgaa	1920
Qy	2161	gtctttgttgcctccatggatttttttttccaaagaatctcagggtaccaaatccggccaggag	2220
Db	1921	gtctttgttgcacatggactttttctcaaaagaatctcaagggtaccaaatcttggcaggaa	1980
Qy	2221	gttttcccgatcaccccaaatccaaatgaggatcatcttttttctaactctctttt	2280
Db	1981	gtcttctcccatatcaccccttaactcaaatgaggatcatcttttttctaactctctttt	2040
Qy	2281	tcaactggtcgggacatctttcgaagggggaagctctcttttttactcttatacttttt	2340
Db	2041	tcaactggcttgggacatcttttgcgaagggggaagctctcttttttactcttatacttttt	2100
Qy	2341	tttttggagtgagctcatcttcttggccaggtctggcctgactctggctcactgcacac	2400
Db	2101	tttttttgagtgagctcatcttcttggccaggtctggcctgactctggctcactgcacac	2160
Qy	2401	ctccacttccctgggttccagcgaaltctcctgcctcagcctcctlaagtagctggattaca	2460
Db	2161	ctccacttccctgggttccagcgaaltctcctgcctcagcctcctlaagtagctggattaca	2220
Qy	2461	ggcgctgcacacacaccagctaaatttatlittagcagagatgggggttctactgtgttg	2520
Db	2221	ggcgctgcacacacaccagctaaatttatttttttttttttttttttttttttttttttt	2280
Qy	2521	ggcagctggtcgtgaactcctcagctcgaagtgcaccacccctcagcctcccaagatg	2580
Db	2281	ggcagctggtcgtgaactcctcagctcgaagtgcaccacccctcagcctcccaagatg	2340
Qy	2581	ctaggattacagggccttttgaactcttttctgagtttttatlgacccctctaattctctt	2640
Db	2341	ctaggattacagggccttttgaactcttttctgagtttttatlgacccctctaattctctt	2400
Qy	2641	accagaaatttatacttccaccagcaactcctgactcttttgaaggaggagcctcagttcta	2700
Db	2401	accagaaatttatacttccaccagcaactcctgactcttttgaaggaggagcctcagttcta	2460
Qy	2701	gtccttggctgctggtgttcattgtgttaggaatgaccacgggctcagtttccccattt	2760
Db	2461	gtccttggctgctggtgttcattgtgttaggaatgaccacgggctcagtttccccattt	2520
Qy	2761	gtataatgggaagcctgtaccaggtcatcttctaagatttctcctgactccagtgagctgg	2820
Db	2521	gtataatgggaagcctgtaccaggtcatcttctaagatttctcctgactccagtgagctgg	2580
Qy	2821	aattctaattgctgctcaggagctgtctccaggatggtgcagatggcttgcggaag	2880
Db	2581	aattctaattgctgctcaggagctgtctccaggatggtgcagatggcttgcggaag	2640
Qy	2881	gagatgggtttggaggccaacaacacctgtctgtcaaatattgccttgcctcttggcagcc	2940
Db	2641	gagatgggtttggaggccaacaacacctgtctgtcaaatattgccttgcctcttggcagcc	2700
Qy	2941	cttgaacttgagtaataacaactccctgaacctcagtttctctcatctgcagaatggga	3000
Db	2701	cttgaacttgagtaataataacaactccctgaacctcagtttctctcatctgcagaatggga	2760
Qy	3001	taattatgtcccaggggataatttagacctgtttcttccaggagggtccccagctcgtt	3060
Db	2761	taattatgtcccaggggataatttagacctgtttcttccaggagggtccccagctcgtt	2820





```

/gene="dJ101A2.2"
/note="match: CDNA5: Em: AK013049 Em: AK011574
match: ESTs: Em: AA233726 Em: T87841 Em: AA756018 Em: AA113277
Em: AA985248 Em: AA682750 Em: AA461487 Em: A1028588
Em: AA659710 Em: A1149891 Em: AA579641 Em: AA516077
Em: AA460132 Em: AA483490 Em: AA046575"
/evidence="not_experimental
/product="dJ28H20.2 (novel protein)"
/complement(1..3110)
/gene="dJ101A2.2"
18..494
/note="match: STS: Em: HS28H20T"
complement(join(84..562,2463..2745))
/gene="dJ101A2.2"
/note="supported by FEENESH and GENSCAN
continues in Em: AL133520 as dJ101A2.2"
/codon_start=1
/evidence="not_experimental
/product="dJ28H20.2 (novel protein)"
/protein_id="CAC00561.1"
/db_xref="GI:9588402"
/db_xref="SPTRMBL:Q9NQE6"
/translation="MAAARATTPADGEPAEALAAARSRSSFLSLGVKQGAE
ARVRFGQRAAVTKHRPKGYRHPALRLGRRTVQEARLLRCRAGISAPVWF
FVDYASNLVWEEIEGCVTVRDYIQSTWETKTPGGLSLARKTIGQVILARMHDEDLIH
GDLTTSNMLLAPPLDQNLIVLIDFGLSFISALPDKGDVLYVLEKAFSLTHPTETVF
EAFKSYSTSKKARPVLKLDVLRGRKRWVG"
867..1145
/note="AlusX repeat: matches 1..298 of consensus"
1251..1353
/note="MIR repeat: matches 50..151 of consensus"
1672..1797
/note="MIR repeat: matches 8..135 of consensus"
1855..2079
/note="MIR repeat: matches 33..262 of consensus"
2155..2248
/note="MIR repeat: matches 33..135 of consensus"
2384..3150
/note="CpG island"
/evidence="not_experimental
2812..2871
/note="f3 copies 20 mer 85% conserved"
/complement(2876..3110)
/gene="dJ101A2.2"
/note="match: STS: Em: G15621"
3865..3962
/note="MER21B repeat: matches 299..391 of consensus"
3963..5655
/note="L1A8 repeat: matches 4532..6270 of consensus"
5669..5895
/note="L1ME repeat: matches 5501..5739 of consensus"
6332..6548
/note="L2 repeat: matches 2524..2746 of consensus"
6626..6915
/note="AlusX repeat: matches 3..290 of consensus"
6918..6980
/note="MER94 repeat: matches 1..62 of consensus"
7023..7322
/note="AluJo repeat: matches 1..312 of consensus"
7323..7417
/note="L2 repeat: matches 2389..2513 of consensus"
7457..7623
/note="MIR repeat: matches 93..262 of consensus"
7864..7934
/note="MLT11 repeat: matches 311..382 of consensus"
8046..8240
/note="MLT1J repeat: matches 9..211 of consensus"
8347..8647
/note="AlusQ repeat: matches 1..300 of consensus"
8685..8793
/note="MLT1J repeat: matches 271..366 of consensus"
9095..9221
/note="MIR repeat: matches 102..242 of consensus"

```

```

repeat_region
9647..9776
/note="L2 repeat: matches 2170..2303 of consensus"
9793..10028
/note="AluDo repeat: matches 74..306 of consensus"
10038..10206
/note="LTR16C repeat: matches 126..316 of consensus"
10631..10690
/note="MLT1D repeat: matches 445..505 of consensus"
10691..10995
/note="AluDb repeat: matches 1..295 of consensus"
10996..11200
/note="MLT1D repeat: matches 248..445 of consensus"
11232..11634
/note="MSTB repeat: matches 2..409 of consensus"
11670..11761
/note="MLT1D repeat: matches 128..192 of consensus"
11762..12062
/note="AlusX repeat: matches 1..293 of consensus"
12063..12180
/note="MLT1D repeat: matches 1..128 of consensus"
complement(12181..12344)
/note="match: GSS: Em: AQ190762"
13011..13129
/note="MLT1J repeat: matches 252..368 of consensus"
13210..13392
/note="MER20 repeat: matches 1..215 of consensus"
13397..13442
/note="MIR repeat: matches 104..146 of consensus"
13660..13783
/note="LTR33 repeat: matches 330..452 of consensus"
13894..14254
/note="L2 repeat: matches 2255..2619 of consensus"
14305..14435
/note="L2 repeat: matches 2617..2748 of consensus"
14446..14525
/note="4 copies 20 mer 76% conserved"
15476..15752
/note="AluJb repeat: matches 2..278 of consensus"
15774..15952
/note="MER58A repeat: matches 42..220 of consensus"
16190..16304
/note="L2 repeat: matches 2657..2744 of consensus"
16305..16599
/note="AluJb repeat: matches 1..296 of consensus"
16600..16635
/note="L2 repeat: matches 2617..2657 of consensus"
16641..16803
/note="L1MA10 repeat: matches 6162..6284 of consensus"
16804..17190
/note="MER57B repeat: matches 1..403 of consensus"
17198..17297
/note="5 copies 20 mer 71% conserved"
17204..17297
/note="47 copies 2 mer ta 70% conserved"
17559..17868
/note="AlusX repeat: matches 1..312 of consensus"
17939..18240
/note="LTR26 repeat: matches 246..553 of consensus"
complement(18333..18408)
/note="match: STS: Em: G42885"
18451..18615
/note="FAM repeat: matches 126..41 of consensus"
19258..19336
/note="MER57-internal repeat: matches 7168..7244 of
consensus"
19343..19485
/note="MER53 repeat: matches 2..180 of consensus"
19721..20190
/note="LTR8 repeat: matches 251..691 of consensus"
20238..20493
/note="LTR8 repeat: matches 1..256 of consensus"
20704..20978

```

Query Match 41.28; Score 1810; DB 9; Length 127418;  
Best Local Similarity 99.94; Pred. No. 0;  
Matches 1910; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2476 accagctcaatttttttagacagatgggggtttcactgtgttgccagcctgctcgtc 2535  
|||||  
Db 47767 ACCAGCTAAATTTTACAGAGATGGGGTTTCACTGTGTGGCCAGCCTGGCTCGTG 47826  
|||||  
Qy 2536 aactcctgagctcaagtgatccacccacactcagcctccacagtgctgtaggattacaggcc 2595  
|||||  
Db 47827 AACTCCTGAGCTCAAGTGATCCACCCACTCAGCCTCCACAGTGTAGGATTACAGGCC 47886  
|||||  
Qy 2596 ttttgactctttttatctgaggtttttatgagaccctcttaattctttacccagaattttat 2655  
|||||  
Db 47887 TTTTGACTCTTTTATCTGAGTTTATTGACCCCTCTAATCTCTTACCCAGAAATTTAT 47946  
|||||  
Qy 2656 ccttcacagcaactcgaactcttgacgggggagcctcagttctagtcctlggtcgtcgtg 2715  
|||||  
Db 47947 CCTTACCAGCAACTCTGACTCTTTGACGGGAGGCCCTCAGTTCTAGTCTCTTGGCTGCTG 48006  
|||||  
Qy 2716 gtgtcaattgctgtggaatgacccagggcctcagtttccccatttgataatgggaagcc 2775  
|||||  
Db 48007 GTGTCAATTTGCTAGGAATGACCCAGGGCCTCAGTTTCCCCATTTGTATAATGGGAAGCC 48066  
|||||  
Qy 2776 tgtaccaggtcatcttaagatttctcctgactccagtgagctggaaattctaaatgctgg 2835  
|||||  
Db 48067 TGTACCAAGGTCAATCTTAAGATTTCTCTGACTCCAGTGAGTGGAAATCTAAATGCTGG 48126  
|||||  
Qy 2836 tctaggagctgtctccaggaagggtcaggaatggctttcggaaaggaagatgggttttgag 2895  
|||||  
Db 48127 TCTAGGAGCTGTCTCCAGAGATGGTGACAGATGGCTTTCCGGANAGGAGATGGCTTTGGAG 48186  
|||||  
Qy 2896 gccacaaacactgctgtcaattatgctttgctcctcttgagcgccttgaacttgagttaa 2955  
|||||  
Db 48187 GCCAAACAACTGCTGTGCAATATTGCTCTTGGCCAGCCCTTGAACCTTGAGTAA 48246  
|||||  
Qy 2956 ataacaactcctcgaacccctcagtttctcctcctcaggaatggggaataattatgtccccgg 3015  
|||||  
Db 48247 ATAACAACCTCCTGAACCTCAGTTTCTCTCATCTGCAGAAATGGGATAATATTGCTCCAGG 48306  
|||||  
Qy 3016 ggtatatttagacctgttctcttccagaggggtccccagctgggtccagggcctgggaaa 3075  
|||||  
Db 48307 GGTATATTTAGACCTGTCTCTTCCAGAGGGTCCCGAGCTGGTCCAGGGCCTGGGAAA 48366  
|||||  
Qy 3076 tttctacttactcctcattccacaggtcctcctcttgagccctgttaaggggtcaggggtgaa 3135  
|||||  
Db 48367 TTTCTACTTCTCTCATTTACCAGGTCTCTCTTTGGACCTGTATAAGSGCTCAGGGTCAA 48426  
|||||  
Qy 3136 tcaagtggggagctgagcaagtagctatgactgagatcactgtaagggagggagactgacaa 3195  
|||||  
Db 48427 TCAGATGGGGAGCTGAGCAAGTAGTATGACTGCAGATCATGTAAAGGAAGGAGCTGACAA 48486  
|||||  
Qy 3196 gaagctccagatgctggggaagtagagagctaaaaatagatcctcaggtcgtgagtgctt 3255  
|||||  
Db 48487 GAAGCTCCACATGCTGGGGAGATGAAGAGCTAAATAGATCTCTAGGTGCTGGATGCTT 48546  
|||||  
Qy 3256 tgtcatcctatcgtgacacatatgggtgctgagagcccccagagactcctggcctctcga 3315  
|||||  
Db 48547 TGTCAATCCATGGTGCACATATGGGTGCTGGCAGAGCCCCCAAGGACTCTGGCCTCTCGA 48606  
|||||  
Qy 3316 gttctcctattcttccattcttagatgcttcccttgatccagtgatgctgctggagctgg 3375  
|||||  
Db 48607 GTTCTCTCTATCTCTCTCCATCTTAGATGCTCTCTCTTGTATCCAGTGTAGTGGAGCTGG 48666  
|||||  
Qy 3376 ctttgcaagctgtgagagctggtgtgacatacttttcaggattttttacaaagtttgtaaac 3435  
|||||  
Db 48667 CTTTGGCAAGCTTGTGAGAGCTGGTGTACATTTTTCAGGATTTTTCAGAGTTGTGTAAC 48726  
|||||  
Qy 3436 acagcattataaaaaataaagttaatttataatttaagtaaaattacattaaaaa 3495  
|||||  
Db 48727 ACAGCCATTATAAAAAATTAATGATTTAAATTAATTAATTAATTAATTAATTAATTA 48786  
|||||  
Qy 3496 aaaaaattataactcaaaattcattacttaattttactacctgttactattatctgtgctt 3555  
|||||

RESULT 5  
AX150151/c  
LOCUS  
DEFINITION Sequence 126 from Patent WO0136685.  
ACCESSION AX150151  
VERSION AX150151.1 GI:14348179  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
AUTHORS Kroes, R.A., Moskalev, J.R. and Yamamoto, H.

Db 48787 AAAAAATTTATTTACTCAAAATTTACTTTAAATTTTACTACCTGTTACTATTATTTCTGTGCTT 48846  
|||||  
Qy 3556 ttgaggtctattctacatagtaactctcttgagagaccttagggagacaccccgcatctct 3615  
|||||  
Db 48847 TTGAGGCTATTCTTACATAGTAACCTTTATGGAGACCTTAGGGGACACACCGGCATCTCT 48906  
|||||  
Qy 3616 tcttgattcccccactcctaatgacatcatcttagtcttggttgctctcaactggtgctg 3675  
|||||  
Db 48907 TCCCTGATTTCCCACTCAATGACATCATGTTAGTCTTTGTTGCTTTAACTGGCTGTGGGA 48966  
|||||  
Qy 3676 gttgttttctatccacaaagatttagagggagactcacatccagggctgtgattcttct 3735  
|||||  
Db 48967 GTGTTTTTGTATCACAAAGATTAGAGAGACTACACATCAGGGCTTGATTTATTGTTGT 49026  
|||||  
Qy 3736 tgaattctgactctcagaacatgctgggtatataaaagtgctcagtaactgcaaatlaaaactttaa 3795  
|||||  
Db 49027 TGATTTTCTAGACTTCAGAACATGCTGGATAAATGTCAGTAATGCAAAATTAACATTTAA 49086  
|||||  
Qy 3796 agtatgtctctgttctgtagcgaatacatctggtgtatagaccccaaaaaaattggagggattatc 3855  
|||||  
Db 49087 AGTATGCTCTTCTGTTAGCTTAATACATGCTGTATAGCACCAAAAAAATGAGGGATTATTC 49146  
|||||  
Qy 3856 ttcagtagttgaaacactgctcctcgtttccagctgacagctgctcaaatcatttaagaag 3915  
|||||  
Db 49147 TTCCAGTACTTGAACACATGCTCATCCGTTTCAGCTGCACAGCTGCTCAATCATTTTAAGAAG 49206  
|||||  
Qy 3916 gagtctgacatcttcttctggttttacttttctcttctcctcactagtgtaaacacaaa 3975  
|||||  
Db 49207 GAGTTCTGACATTTCTATTTGTTTACTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49266  
|||||  
Qy 3976 tttoaacagcattctcgtccgaacctatccacctctcctcagctgctgtagctgtagctat 4035  
|||||  
Db 49267 TTTCAACAGCATTCATCCGAACCTATACCCATTTCTTTCAGTGCCTAGCTGTACAGTTAT 49326  
|||||  
Qy 4036 cagggaatttttattctgtagtctaaattttgtcctaaatcagggccaaatcgagctgtagctg 4095  
|||||  
Db 49327 CAGGATTTTATTTGTTAGTCTTAATTTTGTCAAAATCATGGCCAAATCGCAGTGTAGTTG 49386  
|||||  
Qy 4096 accttggatcaaaagcttttggcaaaaaaaaataataaacaanaattctctgtaagaactcaa 4155  
|||||  
Db 49387 ACTTTGGATACAGGTTTGGCAAAAAAATAATTAACAAATATTCTGTGAAGATCAA 49446  
|||||  
Qy 4156 ttggctatattggaattgagataaagaatatttacaataaagaatatttacaataaagaag 4215  
|||||  
Db 49447 TTGCTATATGGAATTTAGGATAAAGAAATATTACAATAAAGAAATATTACAATAAAGAG 49506  
|||||  
Qy 4216 ttattattatttgaattgtagtgcaacaacataacacctttattctctgttaaaaattata 4275  
|||||  
Db 49507 TTTATTTATTTTGTAGTGTGTCACAAACAAACATACCTTTTATCTCTGTAAATTTATA 49566  
|||||  
Qy 4276 cacacaaaaattacaacaaagattctgtaagaatttaattggctatatggaaatttaggatag 4335  
|||||  
Db 49567 CACACAAAAATTAACAAAAAGATTTCTGAAGAAATTAATTTGGCTATATGGAATTTAGATAG 49626  
|||||  
Qy 4336 aatattacaataaagagatttacaataaagaatttggattatttggtaa 4387  
|||||  
Db 49627 AATATTTACAATAAAGAGATTATTACAATAAAGAGTTTGTATTATTATTGTTAAA 49678  
|||||



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: TRAL Plate: 27 Row: c Column: 21  
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

## FEATURES

Location/Qualifiers

1. .1520

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="MGC:19694 IMAGE:3506460"

/tissue\_type="Kidney, renal cell adenocarcinoma"

/clone\_lib="NIH\_MGC\_14"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

29. .922

/codon\_start=1

/product="Unknown (protein for MGC:19694)"

/protein\_id="AAH10033.1"

/db\_xref="GI:14603130"

/translation="MDAEGIALLPVPTLAALVDSWLRDPCGLNYAALVSGAGPSQA

ALMAKSPGVLGAGPFDIAFTQLNCOVSMFLPEGSKLPVVARVAEVRGPALCLLLGER

VALNTLARGSGIAGAAAFAARGAGTGHVAGTRKTTCGFRIVKYLIVGGAASH

RYDLGLVMKDNHVAAGGVEKAVRAARQADFPALKEVECCSSLOEAVQAEAGADL

VLLNFRPEELHPATVILKAQPPSAVEASGGITLDNLPPQCGPHIDVISMGMLTQAA

PALDFSLLKFAKEVAPVKTH"

BASE COUNT 316 a 434 c 491 g 279 t

## ORIGIN

Query Match 2.3%; Score 99; DB 9; Length 1520;

Best Local Similarity 100.0%; Pred. No. 1e-43;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2389 gctcactcaacctcacttcctgggttcagcgattctcctgcctcagctcctaagta 2448

Db 1444 GCTCACTGCAACCTCCACTTCTGGGTTCAGCGGATTCTCTGCGCTCAGCCCTCTAAGTA 1385.

Qy 2449 gctgggattacagcgcgctgcccacacacccagctcaatt 2487

Db 1384 GCTGGGATTACAGGCGCTGCCACACACCCAGCTAATT 1346

## RESULT

BC005060/c

LOCUS

DEFINITION

Homo sapiens

1524 bp mRNA

PRI

12-JUL-2001

Homo sapiens

Similar to quinolinate phosphoribosyltransferase

(nicotinate-nucleotide pyrophosphorylase (carboxylating)), clone

MGC:12951 IMAGE:2960170, mRNA, complete cds.

ACCESSION

BC005060

VERSION

BC005060.1

KEYWORDS

GI:13477196

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1524)

Strausberg, R.

Direct Submission

Submitted (26-MAR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>

Project Information  
Center Project Name: 600095  
Center clone name: RPCI-11\_50412

#### Summary Statistics

Consensus quality: 172472 bases at least Q40  
Consensus quality: 174610 bases at least Q30  
Consensus quality: 175032 bases at least Q20  
Estimated insert size: 190110; agarose-fp estimation  
Quality coverage: 10.66 in Q20 bases; agarose-fp estimation  
Quality coverage: 11.57 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.

\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1 22409: contig of 22409 bp in length  
22410 22509: gap of unknown length  
22510 31954: contig of 9445 bp in length  
31955 32054: gap of unknown length  
32055 116297: contig of 84243 bp in length  
116298 116397: gap of unknown length  
116398 145662: contig of 29165 bp in length  
145663 145662: gap of unknown length  
145663 175599: contig of 29937 bp in length.

#### FEATURES

Location/Qualifiers  
1. 175599  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-50412"  
/clone\_lib="RPCI human BAC library 11"  
BASE COUNT 43184 a 44157 c 44428 g 43430 t 400 others  
ORIGIN

Query Match 2.3%; Score 99; DB 2; Length 175599;  
Best Local Similarity 100.0%; Pred. No. 1.7e-43;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2389 gctcactgcaacctcactctcctgggttcagcgattctcctgctcagcctccttaagta 2448  
|||||  
Db 6908 GCTCACTGCAACCTCCACTCTCTGGGTTCAAGCGATTCTCTGCGCTCAGCTCTTAAGTA 6849

QY 2449 gctgggattacaggcgctgcccaccacacccagcttaatt 2487  
|||||  
Db 6848 GCTGGGATTACAGCGCGGTGCCACACACCCAGCTTAATT 6810

#### RESULT 10

AC023831  
LOCUS AC023831 181705 bp DNA HTG 25-APR-2001  
DEFINITION HOMO sapiens chromosome 16 clone RP11-74E23, WORKING DRAFT  
SEQUENCE, 8 unordered pieces.

ACCESSION AC023831  
VERSION AC023831.7 GI:13786376

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 181705)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16

#### JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Unpublished  
2 (bases 1 to 181705)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94596, USA  
On Apr 25, 2001 this sequence version replaced gi:11178070.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>

Project Information  
Center Project Name: 0  
Center clone name: RPCI-11\_74E23

#### Summary Statistics

Consensus quality: 169715 bases at least Q40  
Consensus quality: 174817 bases at least Q30  
Consensus quality: 176639 bases at least Q20  
Estimated insert size: 217000; agarose-fp estimation  
Estimated insert size: 181005; sum-of-contigs estimation  
Quality coverage: 9.4 in Q20 bases; agarose-fp estimation  
Quality coverage: 11.27 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1212: contig of 1212 bp in length  
1213 1312: gap of unknown length  
1313 2398: contig of 1086 bp in length  
2399 2498: gap of unknown length  
2499 3678: contig of 1180 bp in length  
3679 3778: gap of unknown length  
3779 7829: contig of 4051 bp in length  
7830 7929: gap of unknown length  
7930 27212: contig of 19283 bp in length  
27213 27312: gap of unknown length  
27313 49715: contig of 22403 bp in length  
49716 49815: gap of unknown length  
49816 88863: contig of 39048 bp in length  
88864 88963: gap of unknown length  
88964 181705: contig of 92742 bp in length.

#### FEATURES

Location/Qualifiers  
1. 181705  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-74E23"  
/clone\_lib="RPCI human BAC library 11"  
BASE COUNT 45562 a 43824 c 45067 g 45113 t 2139 others  
ORIGIN

#### Query Match

2.3%; Score 99; DB 2; Length 181705;  
Best Local Similarity 100.0%; Pred. No. 1.7e-43;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2389 gctcactgcaacctcactctcctgggttcagcgattctcctgctcagcctccttaagta 2448  
|||||

Db 42813 GCTCACTGCAACCTCCACTCTCTGGGTTCAAGCGATTCTCTGCGCTCAGCTCTTAAGTA 42872

QY 2449 gctgggattacaggcgctgcccaccacacccagcttaatt 2487  
|||||

Db 42873 GCTGGGATTACAGCGCGGTGCCACACACCCAGCTTAATT 42911

#### RESULT 11

AC023105



LOCUS AC023105 70876 bp DNA PRI 09-MAY-2001  
 DEFINITION Homo sapiens BAC clone GSI-13M19 from 7, complete sequence.  
 ACCESSION AC023105  
 VERSION AC023105.7 GI:11415202  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 70876)  
 AUTHORS Sulston, J.E. and Waterston, R.  
 JOURNAL Toward a complete human genome sequence  
 MEDLINE Genome Res. 8 (11), 1097-1108 (1998)  
 99063792  
 REFERENCE 2 (bases 1 to 70876)  
 AUTHORS Mulvane, E., Nguyen, C., Strommatt, C. and Waligorski, J.  
 JOURNAL The sequence of Homo sapiens BAC clone GSI-13M19  
 TITLE Unpublished  
 REFERENCE 3 (bases 1 to 70876)  
 AUTHORS Waterston, R.H.  
 JOURNAL Direct Submission  
 TITLE Submitted (08-FEB-2000) Genome Sequencing Center, Washington  
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 REFERENCE 4 (bases 1 to 70876)  
 AUTHORS Waterston, R.H.  
 JOURNAL Direct Submission  
 TITLE Submitted (29-NOV-2000) Genome Sequencing Center, Washington  
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 REFERENCE 5 (bases 1 to 70876)  
 AUTHORS Waterston, R.  
 JOURNAL Direct Submission  
 TITLE Submitted (09-MAY-2001) Department of Genetics, Washington  
 JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 COMMENT On Nov 29, 2000 this sequence version replaced gi:9838027.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@genome.wustl.edu](mailto:sapiens@genome.wustl.edu)  
 ----- Summary Statistics  
 ----- Center project name: H\_GS013M19  
 -----  
 NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.  
 This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.  
 MAPPING INFORMATION:  
 The sequence of this clone was established as part of a mapping and  
 sequencing collaboration between the NIGRI Chromosome 7 Mapping  
 Project (Eric D. Green, Director), John D. McPherson in the  
 Department of Genetics (Washington University), and the Washington  
 University Genome Sequencing Center. For additional information  
 about the map position of this sequence, see  
<http://www.nhgri.nih.gov/DIR/OTB/CHR7>, send  
<mailto:egreen@ngri.nih.gov>, or see <http://genome.wustl.edu/gsc>  
 SOURCE INFORMATION:  
 This clone is from the first BAC library from Genome Systems, Inc.  
 (<http://www.genomesystems.com>).  
 Cell line: lymphoblastoid  
 Haplotypes: two

VECTOR: pBelOBAC  
 SELECTION: chloramphenicol  
 NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the right is GSI-96J14, 200 bp overlap.  
 Actual start of this clone is at base position 1 of GSI-13M19;  
 actual end is at base position 63216 of GSI-96J14.  
 The clone GSI-13M19 contains a transposon which is not part of the  
 submitted sequence.  
 FEATURES  
 source 1..70876  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7"  
 /clone="GSI-13M19"  
 /clone\_lib="GSBAC1"  
 134..433  
 /rpt\_family="Alu"  
 repeat\_region 529..810  
 /rpt\_family="Alu"  
 repeat\_region 1312..1585  
 /rpt\_family="Alu"  
 repeat\_region 1589..1884  
 /rpt\_family="Alu"  
 repeat\_region 1995..2125  
 /rpt\_family="Alu"  
 misc\_feature 2322..2407  
 /note="similar to EST BE796113 (NID:g10217311)"  
 misc\_feature 2325..2434  
 /note="similar to EST AI663572 (NID:g4767155) u198e02.y1"  
 misc\_feature 2334..2407  
 /note="similar to EST AI196214 (NID:g3748820) u170d05.y1"  
 misc\_feature 2337..2404  
 /note="similar to EST AAS89188 (NID:g2405089) vk08f11.r1"  
 misc\_feature 2338..2404  
 /note="similar to EST AW763036 (NID:g7694906) ur68c05.y1"  
 misc\_feature 2338..2404  
 /note="similar to EST BF233698 (NID:g11144387)"  
 misc\_feature 2342..2404  
 /note="similar to EST AW761967 (NID:g7693882) ur52a03.y1"  
 misc\_feature 2342..2404  
 /note="similar to EST BF233846 (NID:g11144691)"  
 misc\_feature 2342..2404  
 /note="similar to EST BF234140 (NID:g11145301)"  
 misc\_feature 2344..2404  
 /note="similar to EST BF164569 (NID:g11044894)"  
 misc\_feature 2345..3115  
 /note="similar to EST BF165458 (NID:g11045810)"  
 misc\_feature 2380..2525  
 /note="similar to EST BE882243 (NID:g10331019)"  
 misc\_feature 2382..2525  
 /note="similar to EST BE545771 (NID:g9774416)"  
 misc\_feature 2382..2525  
 /note="similar to EST BE788953 (NID:g10210151)"  
 misc\_feature 2382..2525  
 /note="similar to EST BE262701 (NID:g9135941)"  
 misc\_feature 2384..2525  
 /note="similar to EST BE737523 (NID:g10151515)"  
 misc\_feature 2385..2525  
 /note="similar to EST BE258933 (NID:g9129429)"  
 misc\_feature 2385..2525  
 /note="similar to EST BF141274 (NID:g10980314)"  
 misc\_feature 2386..2525  
 /note="similar to EST BE514331 (NID:g9721543)"  
 misc\_feature 2389..2525  
 /note="similar to EST BE798937 (NID:g10220135)"  
 misc\_feature 2391..2525  
 /note="similar to EST BF036990 (NID:g10745154)"  
 misc\_feature 2392..2525  
 /note="similar to EST BE270710 (NID:g9144352)"  
 misc\_feature 2392..2525  
 /note="similar to EST BF203170 (NID:g11096809)"



\* 46605 51108: contig of 4504 bp in length  
 \* 51109 51208: gap of unknown length  
 \* 55209 55594: contig of 4386 bp in length  
 \* 55595 55594: gap of unknown length  
 \* 55695 61762: contig of 6068 bp in length  
 \* 61763 61862: gap of unknown length  
 \* 61863 69971: contig of 8109 bp in length  
 \* 69972 70071: gap of unknown length  
 \* 70072 79254: contig of 9183 bp in length  
 \* 79255 79354: gap of unknown length  
 \* 79355 94310: contig of 14856 bp in length  
 \* 94311 94310: gap of unknown length  
 \* 111262 111362: contig of 16952 bp in length  
 \* 111363 132866: contig of 21504 bp in length  
 \* 132867 132966: gap of unknown length  
 \* 132967 159749: contig of 26783 bp in length  
 \* 159750 159849: gap of unknown length  
 \* 159850 185926: contig of 26077 bp in length.

# FEATURES

source  
 1..185926  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP11-487E1"  
 1..1472  
 /note="assembly\_name:Contig16"  
 1573..2852  
 /note="assembly\_name:Contig17"  
 2953..4741  
 /note="assembly\_name:Contig18"  
 4842..8040  
 /note="assembly\_name:Contig19"  
 8141..11050  
 /note="assembly\_name:Contig20"  
 11151..14348  
 /note="assembly\_name:Contig21"  
 14449..18009  
 /note="assembly\_name:Contig22"  
 clone\_end:T7  
 vector\_side:left  
 18110..22346  
 /note="assembly\_name:Contig23"  
 22447..27263  
 /note="assembly\_name:Contig24"  
 27364..31414  
 /note="assembly\_name:Contig25"  
 31515..36252  
 /note="assembly\_name:Contig26"  
 36353..41629  
 /note="assembly\_name:Contig27"  
 41730..46504  
 /note="assembly\_name:Contig28"  
 46605..51108  
 /note="assembly\_name:Contig29"  
 51209..55594  
 /note="assembly\_name:Contig30"  
 55695..61762  
 /note="assembly\_name:Contig31"  
 61863..69971  
 /note="assembly\_name:Contig32"  
 70072..79254  
 /note="assembly\_name:Contig33"  
 79355..94210  
 /note="assembly\_name:Contig34"  
 94311..111262  
 /note="assembly\_name:Contig35"  
 111363..132866  
 /note="assembly\_name:Contig36"  
 clone\_end:SP6  
 vector\_side:right  
 132967..159749  
 /note="assembly\_name:Contig37"

misc\_feature 159850..185926  
 /note="assembly\_name:Contig38"  
 BASE COUNT 57022 a 36205 c 36048 g 54445 t 2206 others  
 ORIGIN  
 Query Match 1.5%; Score 64; DB 2: Length 185926;  
 Best Local Similarity 100.0%; Pred. NO. 8.1e-24;  
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2380 ctgatttggctcactgcacacctccattctctctgggttcaggatctctctgcctcagcc 2439  
 |||||  
 DB 82937 CTGATCTGGCTGACATGCAACCTCCACTTCTCTGGGTTCAGACGATTCTCTGCTCAGCC 82996  
 |||||  
 QY 2440 tctc 2443  
 |||||  
 DB 82997 TCCT 83000  
 |||||  
 RESULT 13  
 AC020960  
 LOCUS 108162 bp DNA HTG 18-JUL-2000  
 DEFINITION Mus musculus clone CT7-240L13, WORKING DRAFT SEQUENCE, 1 ordered  
 pieces.  
 ACCESSION AC020960  
 VERSION AC020960.2 GI:9256383  
 KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 108162)  
 TITLE Sequencing of Mouse  
 JOURNAL Unpublished  
 2 (bases 1 to 108162)  
 DOE Joint Genome Institute.  
 Direct Submission  
 Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 On Jul 18, 2000 this sequence version replaced gl:6691270.  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: http://www.jgi.doe.gov  
 -----  
 Project Information  
 Center Project Name: 1045618  
 Center clone name: RG-MBAC\_240L13  
 -----  
 Summary Statistics  
 Consensus quality: 103750 bases at least Q40  
 Consensus quality: 106923 bases at least Q30  
 Consensus quality: 107282 bases at least Q20  
 Estimated insert size: 130000; pulse field gel estimation  
 Quality coverage: 6.53 in Q20 bases; sum-of-contigs estimation  
 Quality coverage: 7.85 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 108162: contig of 108162 bp in length.  
 Location/Qualifiers  
 1..108162  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="CT7-240L13"



```
Center Project Name: 704688
Center clone name: C17B-H1_220003
-----
Summary Statistics
Consensus quality: 116661 bases at least Q40
Consensus quality: 120243 bases at least Q30
Consensus quality: 121387 bases at least Q20
Estimated insert size: 125000; pulse field gel estimation
Estimated insert size: 121972; sum-of-contigs estimation
Quality coverage: 8.51 in Q20 bases; pulse field gel estimation
Quality coverage: 8.72 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
36662: contig of 36662 bp in length
36663
36762: gap of unknown length
36763
45137: contig of 8375 bp in length
45138
45237: gap of unknown length
45238
54176: contig of 8939 bp in length
54177
54276: gap of unknown length
54277
85719: contig of 31443 bp in length
85720
88481: contig of 2662 bp in length
88482
88581: gap of unknown length
88582
93330: contig of 4749 bp in length
93331
96495: gap of unknown length
96496
96595: gap of unknown length
96596
119670: contig of 23075 bp in length
119671
119771
122772: contig of 3002 bp in length.
FEATURES
source
1..122772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-220003"
/clone_lib="CalTech human BAC library D"
BASE COUNT 39486 a 25066 c 23097 g 34312 t 811 others
ORIGIN

Query Match 1.4%; Score 63; DB 2; Length 122772;
Best Local Similarity 100.0%; Pred. No. 2.8e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2381 tgatcttggtcactgcacactccacttcctcgggttcgaaggattctctgcctcagcct 2440
|||||
Db 120510 TGATCTTGGGCTCACTGCAACCTCCACTTCCTGGGTCAAGCGATTCTCCTGCTCAGCCT 120569

Qy 2441 cct 2443
|||
Db 120570 CCT 120572

Search completed: March 15, 2002, 07:03:55
Job time: 10176 sec
```

**THIS PAGE BLANK (USPTO)**